Objective Bayesian Analysis for a Truncated Model

Haiying Wang

Department of Statistics, University of Missouri, Columbia, Missouri 65211, USA

Dongchu Sun

Department of Statistics, University of Missouri, Columbia, Missouri 65211, USA and School of Finance and Statistics, East China Normal University, Shanghai 200241, China

Abstract

In this paper, the reference prior is developed for a truncated model with boundaries of support as two functions of an unknown parameter. It generalizes the result obtained in a recent paper by Berger *et al.* (2009), in which a rigorous definition of reference priors was proposed and the prior for a uniform distribution with parameter-dependent support was derived. The assumption on the order of the derivatives of these two boundary functions, required by Berger *et al.* (2009), is removed. In addition, we obtain the frequentist asymptotic distribution of Bayes estimators under the squared error loss function. Comparisons of Bayesian approach with frequentist approach are drawn in two examples in detail. Both theoretical and numerical results indicate that Bayesian approach, especially under the reference prior, is preferable.

Keywords: Asymptotic, Bayes estimator, Non-regular, Reference prior, Truncated model

1. Introduction

Reference priors, initially introduced by Bernardo (1979), fulfill the purpose of objective Bayesian analysis through maximization of the missing information about the parameter (c.f. Berger *et al.*, 2009). Since the prior information is maximally dominated by the data, reference priors usually have appealing frequentist properties in statistical inference such as point and interval estimation.

Bernardo (1979) originally introduced the concept of reference priors by defining them as priors that maximize the missing information for unknown parameters (c.f. Berger *et al.*, 2009). But the derivation was informal and the explicit expression obtained reduced to Jeffreys' prior (Jeffreys, 1946, 1961) for continuous one parameter models under posterior asymptotic normality. Berger and Bernardo (1989) derived the reference for a normal model when the parameter of interest is a product of two means. Ye and Berger (1991) found the reference prior for a exponential regression model. Berger and Bernardo (1992a,b) developed an sequential reference process for multi-parameter problems, but their consideration focused on continuous

Email addresses: `hwzq7@mail.missouri.edu` (Haiying Wang), `sund@missouri.edu` (Dongchu Sun)

regular problems under asymptotic posterior normality. This reference process was applied to the balanced variance components problem in Berger and Bernardo (1992c). Yang and Berger (1994) investigated the reference prior for estimating the covariance matrix of a multivariate normal distribution. Sun and Ye (1995) generalized Berger and Bernardo (1989)'s result to estimate a product of n $(n > 2)$ normal means. Sun and Berger (1998) studied reference priors based on partial information.

The aforementioned work focused on developing reference priors for regular cases in the sense that the Fisher information exists and the posterior is asymptotically normal. For non-regular cases, reference priors are considerably more challenging to derive and many problems remain to be explored. Ghosal and Samanta (1997) derived reference priors for non-regular models with parameter-dependent support. Ghosal (1997) developed reference priors for multi-parameter densities with discontinuities depending on a single parameter. More related work can be found in Bernardo (2005) and the references therein.

This paper is closely related to Berger *et al.* (2009) who gave a rigorous definition of reference priors. Consider the following truncated model

$$f(x \mid \theta) = \frac{g(x)}{G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)} I_{[a_1(\theta),\, a_2(\theta)]}(x), \quad \theta \in (\theta_L,\, \infty), \tag{1}$$

where θ is an unknown parameter and θ_L is the lower bound of the parameter space; $I_A(\cdot)$ is the indicator function of a set A; $G(\cdot)$ is a strictly increasing and continuously differentiable function so $g(\cdot) = G'(\cdot) > 0$; $a_1(\cdot)$ and $a_2(\cdot)$ are continuously differentiable and strictly increasing functions. Usually, $g(\cdot)$ is a density function on $\mathbb{R}$ or $\mathbb{R}^+$ so that $f(x \mid \theta)$ is the truncation of the density on $[a_1(\theta),\ a_2(\theta)]$.

Model (1) is not a regular model since the support of the distribution depends on an unknown parameter and the Fisher information does not exist. This model is also different from the one in Ghosal and Samanta (1997) in which it is required that one of $a_1(\cdot)$ and $a_2(\cdot)$ is increasing while the other is decreasing. When both $a_1(\cdot)$ and $a_2(\cdot)$ are increasing, Berger *et al.* (2009) derived the reference prior of θ for a uniform distribution, i.e., $g(\cdot) \equiv 1$. Furthermore, the following condition on the derivatives of $a_1(\cdot)$ and $a_2(\cdot)$ is necessary in their theorem,

$$a_1'(\theta) < a_2'(\theta), \quad \theta \in (\theta_L, \infty). \tag{2}$$

In this paper, the result in Berger *et al.* (2009) is generalized. The reference prior of θ for model (1) is developed for any continuous function $g(\cdot) > 0$, and the condition in (2) is waived.

In non-regular estimation problems, Bayesian approach often has superior properties to frequentist approach (see Hill, 1963; Hirano and Porter, 2003; Hall and Wang, 2005). Hirano and Porter (2003) proved that for models with parameter-dependent support, the maximum likelihood estimator (MLE) is generally inefficient, whereas Bayes estimators are efficient in term of the local asymptotic minimax criterion for conventional loss functions. For the model of interest in this paper, we obtain the asymptotic distribution of Bayes estimators under the squared error loss function, which shows that Bayes estimators are asymptotically unbiased. On the other hand, the MLE is usually asymptotically biased, and some time it is not unique. Two examples are elaborated to show the performance of Bayes estimators with reference priors. Systematic simulation studies are carried out to investigate the finite sample performance of Bayesian approaches compared with that of the frequentist approach. Both theoretical and numerical results indicate that Bayes estimators under reference priors have desirable frequentist properties and they are better strategies than the MLE for the model considered in this paper.

The rest of the paper is organized as follows. In Section 2, we present the main results. In Section 3, two examples are discussed in detail and numerical results are provided. Section 4 summarizes and technical details are given in the two appendices.

2. Main Results

This section gives the main results of the paper, including the reference prior for model (1) and the sampling properties of Bayes estimators under the squared error loss function.

2.1. The Reference Prior

Before stating the result, we need to introduce some notations. Let

$$\lambda_i = \frac{G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)}{g\left(a_i(\theta)\right) a_i'(\theta)}, i = 1, 2,$$

and define $\lambda = \lambda_1\lambda_2/(\lambda_1 - \lambda_2)$ if $\lambda_1 \neq \lambda_2$.

Theorem 1. *The reference prior of θ for model* (1) *is*

$$\pi_{\mathrm{R}}(\theta) = \begin{cases} \frac{1}{|\lambda|} \exp\left\{1 + b_1 \vee b_2 + \frac{1}{b_1 - b_2}\left[b_1\psi\left(\frac{1}{b_1}\right) - b_2\psi\left(\frac{1}{b_2}\right)\right]\right\}, & \text{if } \lambda_1 \neq \lambda_2, \\ \frac{1}{\lambda_1}, & \text{if } \lambda_1 = \lambda_2, \end{cases} \tag{3}$$

where $b_1 \vee b_2 = max(b_1, b_2)$, $b_i = \lambda_i/|\lambda|, i = 1, 2$ when $\lambda_1 \neq \lambda_2$, and $\psi(x) = \mathrm{d}\log(\int_0^\infty e^{-t}t^{x-1}\mathrm{d}t)/\mathrm{d}x$ is the digamma function.

Proof. In Appendix A. □

Remark 1. *(a) If $\lambda_1 \equiv \lambda_2$, or equivalently $G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right) \equiv C_g$, a constant, then*

$$\pi_{\mathrm{R}}(\theta) \propto \frac{1}{\lambda_1} \propto g\big(a_1(\theta)\big)a_1'(\theta).$$

Define $y = G(x)/C_g$. Then model (1) *can be transformed to a location family with location parameter $\mu = \{G(a_2(\theta)) + G(a_1(\theta))\}/[2\{G(a_2(\theta)) - G(a_1(\theta))\}]$. The expectation of y is μ, and the reference prior of μ is the constant prior.*

(b) If $\lambda_1 = \lambda_2$ only on a set of θ with Lebesgue measure 0, then

$$\pi_{\mathrm{R}}(\theta) = \frac{1}{|\lambda|} \exp\left\{b_1 \vee b_2 + \frac{1}{b_1 - b_2}\left[b_1\psi\left(\frac{1}{b_1}\right) - b_2\psi\left(\frac{1}{b_2}\right)\right]\right\}. \tag{4}$$

(c) *If* $g(\cdot) \equiv 1$ *and the condition in* (2) *holds, then the prior in* (3) *reduces to the prior obtained in* Berger et al. *(2009). The transformation* $y = G(x)$ *can be utilized to broaden the result in Berger* et al. *(2009). With this transformation,* y *is a uniform distribution with two boundaries of support as another two functions of* θ*. Then Berger et al (2009)'s result can be applied to find the reference prior for* θ*. However the generalization can only be accomplished on* $G(\cdot)$ *such that* $\lambda_1 > \lambda_2$ *for all* θ*, because the condition in* (2) *is required in their result.*

(d) *For* $a_1(\cdot)$ *and* $a_2(\cdot)$ *being both decreasing, the reference prior can be found by using transformation* $y = -x$*.*

2.2. Properties of Bayes Estimators

Suppose a simple random sample $\{x_1, ..., x_n\}$ is taken from model (1). When the sample size is finite, the MLE based on the observed sample is rather complicated, and it may not be unique. A Bayesian estimator, on the other hand, always has explicit expression even though it involves integration. Given a prior $\pi(\theta)$, the Bayes estimator is the posterior mean under the squared error loss function, which takes the form in (5) for model (1).

$$\hat{\theta}_{\mathrm{B}} = \int_{a_2^{-1}(t_n)}^{a_1^{-1}(t_1)} \frac{s\pi(s)}{\{G(a_2(s)) - G(a_1(s))\}^n} \mathrm{d}s \Big/ \int_{a_2^{-1}(t_n)}^{a_1^{-1}(t_1)} \frac{\pi(s)}{\{G(a_2(s)) - G(a_1(s))\}^n} \mathrm{d}s, \tag{5}$$

where $t_1 = x_{(1)}$ and $t_n = x_{(n)}$ are the smallest and largest observations, and $a_1^{-1}(\cdot)$ and $a_2^{-1}(\cdot)$ are the inverse functions of $a_1(\cdot)$ and $a_2(\cdot)$, respectively.

For fixed θ, if $\lambda_1 \neq \lambda_2$, the MLE has the following asymptotic distribution,

$$n\left(\hat{\theta}_{\mathrm{M}} - \theta\right) \xrightarrow{d} \begin{cases} \lambda_1 Z_1, & \text{if } \lambda_1 < \lambda_2, \\ -\lambda_2 Z_2, & \text{if } \lambda_1 > \lambda_2, \end{cases} \tag{6}$$

as $n \to \infty$, where Z_1 and Z_2 are iid $Exp(1)$. So the asymptotic bias is either λ_1 or $-\lambda_2$ instead of 0, meaning that the MLE is always asymptotically biased. If $\lambda_1 = \lambda_2$, the expression of the MLE varies and, in general, a closed form is not available even for a large sample size.

Next theorem gives the asymptotic properties of the Bayes estimator in (5) whose proof is given in Appendix B.

Theorem 2. *For fixed* θ*, the Bayes estimator in* (5) *has the following convergence properties.*

$$n\left(\hat{\theta}_B - \theta\right) \xrightarrow{d} \Lambda \text{ as } n \to \infty,$$

where

$$\Lambda = \begin{cases} \frac{(\lambda - \lambda_2 Z_2)e^{\frac{\lambda_2 Z_2}{\lambda}} - (\lambda + \lambda_1 Z_1)e^{-\frac{\lambda_1 Z_1}{\lambda}}}{e^{\frac{\lambda_2 Z_2}{\lambda}} - e^{\frac{-\lambda_1 Z_1}{\lambda}}}, & \text{if } \lambda_1 \neq \lambda_2. \\ \lambda_1 \frac{Z_1 - Z_2}{2}, & \text{if } \lambda_1 = \lambda_2. \end{cases} \tag{7}$$

Moreover, the Bayes estimator is asymptotically unbiased, i.e. $\mathbf{E}\Lambda = 0$*.*

Remark 2. *(a) The distribution of Λ holds whenever $a_1(\cdot)$ and $a_2(\cdot)$ are monotonic and differentiable.*

(b) If $\lambda_1 = \lambda_2$, the distribution of Λ is symmetric, i.e., $\Lambda \stackrel{d}{=} -\Lambda$. If $\lambda_1 \neq \lambda_2$, the distribution of Λ is asymmetric, but still $\mathbf{E}\Lambda = 0$. The asymmetry of the distribution of Λ is easily seen when λ_1 is much larger than λ_2, because the distribution of Λ is close to that of $\lambda - \lambda_2 Z_2$ in this case.

(c) As in regular cases, the asymptotic distribution Λ is independent of the prior.

3. Examples

Here two examples are given to the illustrate the properties of Bayes estimators and to compare them with that of the frequentist estimators. Numerical studies are carried out to evaluate the finite sample performance and/or the asymptotic properties of both Bayes estimators and the MLE.

3.1. Example 1

Let $\{x_1, ..., x_n\}$ be a simple random sample taken from the density

$$f(x \mid \theta) = \frac{1}{\log(2) x \log(x)} I_{[\theta, \theta^2]}(x). \tag{8}$$

This is a special case of model (1), where $g(x) = 1/\{x \log(x)\}$, $G(x) = \log(\log(x))$, $a_1(\theta) = \theta$ and $a_2(\theta) = \theta^2$ for $\theta \in (1, \infty)$. Note that $g(\cdot)$ is not a density function here.

3.1.1. Theoretical Comparisons

The likelihood function of θ based on the sample is

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\log(2) x_i \log(x_i)} I_{[\sqrt{t_n}, t_1]}(\theta) \propto I_{[\sqrt{t_n}, t_1]}(\theta).$$

The MLE. Clearly, any $\hat{\theta}$ is a MLE if $\sqrt{t_n} \leq \hat{\theta} \leq t_1$. When MLE is not unique, the problem of choosing an estimator is challenging. For simplicity, we focus on the class of MLE that can be written as $\hat{\theta} = (1 - w)t_1 + w\sqrt{t_n}$ for some constant $w \in [0, 1]$. It can be shown that

$$n(\hat{\theta} - \theta) \xrightarrow{d} \lambda_1 \big[(1 - w)Z_1 - wZ_2\big],$$

where $\lambda_1 = \log(2)\theta \log(\theta)$ and Z_1 and Z_2 are iid $Exp(1)$. When $w = 0.5$, the asymptotic variance is minimized and the asymptotic bias is 0. So the asymptotically optimal MLE is

$$\hat{\theta}_{\text{M}}^{\text{O}} = \frac{t_1 + \sqrt{t_n}}{2}. \tag{9}$$

The shortest confidence interval of θ with nominal coverage probability α is

$$\text{CI}_{\text{M}}^{\text{O}} = \Big[\hat{\theta}_{\text{M}}^{\text{O}} + \frac{\log(1 - \alpha) \log(2) \hat{\theta}_{\text{M}}^{\text{O}} \log(\hat{\theta}_{\text{M}}^{\text{O}})}{2n}, \quad \hat{\theta}_{\text{M}}^{\text{O}} - \frac{\log(1 - \alpha) \log(2) \hat{\theta}_{\text{M}}^{\text{O}} \log(\hat{\theta}_{\text{M}}^{\text{O}})}{2n}\Big]. \tag{10}$$

Bayesian Estimation under the Constant Prior. Interestingly, the optimal MLE in (9) is also a Bayes estimator under the constant prior. Since the posterior under the constant prior is a uniform distribution on

$[\sqrt{t_n}, t_1]$, all the Bayesian credible intervals with the same nominal level have equal length. The equal-tailed Bayesian credible interval of θ under the constant prior is

$$\mathrm{CI}_{\mathrm{B}}^{\mathrm{C}} = \left[\frac{t_1 + \sqrt{t_n}}{2} - \alpha\frac{t_1 - \sqrt{t_n}}{2}, \quad \frac{t_1 + \sqrt{t_n}}{2} + \alpha\frac{t_1 - \sqrt{t_n}}{2}\right]. \tag{11}$$

Bayesian Estimation under the Reference Prior. By Theorem 1, the reference prior of θ is

$$\pi_{\mathrm{R}}(\theta) \propto \frac{1}{\theta \log(\theta)}.$$

The corresponding posterior density of θ given (t_1, t_n) is

$$\pi_{\mathrm{R}}\left(\theta \mid t_1, \sqrt{t_n}\right) = \frac{1}{\log(\log(t_1)) - \log(\log(\sqrt{t_n}))} \frac{1}{\theta \log(\theta)} I_{[\sqrt{t_n}, t_1]}(\theta). \tag{12}$$

The posterior mean of θ given the sufficient statistics (t_1, t_n) is

$$\hat{\theta}_{\mathrm{B}}^{\mathrm{R}} = \frac{\int_{\sqrt{t_n}}^{t_1} \frac{1}{\log(\theta)} \mathrm{d}\theta}{\log(\log(t_1)) - \log(\log(\sqrt{t_n}))}. \tag{13}$$

Since the posterior density in (12) is not a constant in the support interval, the shortest Bayesian credible interval of θ with nominal coverage level α is derived as

$$\mathrm{CI}_{\mathrm{B}}^{\mathrm{R}} = \left[\sqrt{t_n}, \quad \sqrt{t_n}^{\left(\frac{\log(t_1)}{\log(\sqrt{t_n})}\right)^{\alpha}}\right], \tag{14}$$

which turns out to be a one sided interval because $\sqrt{t_n}$ is the lower bound of the possible values of θ. In addition, let $u = \log\{\log(x)/\log(\theta)\}/\log(2)$, and then $\Pr(\theta \in \mathrm{CI}_{\mathrm{B}}^{\mathrm{R}} \mid \theta)=\mathrm{P}\left\{\frac{1-u_{(n)}}{1-u_{(n)}+u_{(1)}} < \alpha\right\}$ where $u_{(1)}$ and $u_{(n)}$ are the smallest and largest observations of a sample of size n from Uniform$[0, 1]$. Using similar technique used in the proof of Theorem 1 and by tedious calculation, it can be shown that this probability α at the rate of $1/n$ as $n \to \infty$. So the reference prior is also the second order matching prior for θ.

By re-parametrization $\mu = \log(\log(\theta))/\log(2) + 1/2$ and transformation $y = \log(\log(x))/\log(2)$, model (8) can be transformed to $f(y \mid \mu) \sim$ Uniform $[\mu - 0.5, \mu + 0.5]$, which belongs to the location family. The reference prior of μ is the constant prior and its corresponding posterior is $\pi\left(\mu \mid y_{(1)}, y_{(n)}\right) \sim$ Uniform $[y_{(n)} - 0.5, y_{(1)} + 0.5]$, where $y_{(1)}$ and $y_{(n)}$ are the smallest and largest observed value of y. If the parameter of interest is μ, then the posterior mean and the minimum risk invariant estimator (MRIE) of μ are both $\hat{\mu} = \{y_{(1)} + y_{(n)}\}/2$. However, if $\hat{\mu}$ is transformed back to have an estimator of θ, it is not the MRIE of θ.

3.1.2. Numerical Comparisons

For point estimators, relative mean square errors (RMSE) are calculated, which are ratios of the MSEs of a given estimator to that of $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$. For interval estimators, frequentist coverage probabilities are compared as well as relative average lengths that are defined similarly to RMSE. All the results are reported in Figures .1 and .2 based on 1000 repetitions of the simulation.

Figure .1 displays the results related to $\hat{\theta}_{\mathrm{M}}^{\mathrm{O}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$. Note that $\hat{\theta}_{\mathrm{M}}^{\mathrm{O}}$ is also a Bayes estimator under the constant prior, so we are also comparing Bayes estimators under different priors. It is seen that $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ invariably dominates $\hat{\theta}^{\mathrm{O}}_{\mathrm{M}}$, and the performance of $\hat{\theta}^{\mathrm{R}}_{\mathrm{B}}$ relative to $\hat{\theta}^{\mathrm{O}}_{\mathrm{M}}$ improves as θ grows or as n decreases. When $n = 20$, $\hat{\theta}^{\mathrm{R}}_{\mathrm{B}}$ and $\hat{\theta}^{\mathrm{O}}_{\mathrm{M}}$ have similar accuracy.

Figure .2 presents the results related to the interval estimation. In part (a)–(c), for the extreme case that $n = 1$, the coverage probabilities of $\mathrm{CI}^{\mathrm{O}}_{\mathrm{M}}$ are far from the nominal coverage levels for large α, whereas $\mathrm{CI}^{\mathrm{R}}_{\mathrm{B}}$ always gives coverage probabilities close to the nominal levels. $\mathrm{CI}^{\mathrm{C}}_{\mathrm{B}}$, although being an inferior alternative to $\mathrm{CI}^{\mathrm{R}}_{\mathrm{B}}$, provides better results than $\mathrm{CI}^{\mathrm{O}}_{\mathrm{M}}$. When n is large (=20,500), $\mathrm{CI}^{\mathrm{O}}_{\mathrm{M}}$ gives reasonable outcomes. Part (d) shows the relative average lengths. $\mathrm{CI}^{\mathrm{R}}_{\mathrm{B}}$ dominates $\mathrm{CI}^{\mathrm{O}}_{\mathrm{M}}$ for small sample sizes ($n = 1, 3$), and it dominates $\mathrm{CI}^{\mathrm{C}}_{\mathrm{B}}$ uniformly. There is a tendency for the relative lengths to decrease as n increases, but even when $n = 500$ $\mathrm{CI}^{\mathrm{R}}_{\mathrm{B}}$ still outperforms $\mathrm{CI}^{\mathrm{O}}_{\mathrm{M}}$ for high confidence levels. The two Bayesian approaches are almost equivalent when $n = 500$.

3.2. Example 2

Suppose a simple random sample $\{x_1, ..., x_n\}$ is drawn from the density below,

$$f(x \mid \theta) = \frac{e^{-x}}{e^{-\theta} - e^{-\theta^2}} I_{[\theta,\theta^2]}(x), \quad \theta > 1. \tag{15}$$

Here $g(x) = e^{-x}$, the density of $Exp(1)$, $G(x) = 1 - e^{-x}$, $a_1(\theta) = \theta$ and $a_2(\theta) = \theta^2$ for $\theta \in (1, \infty)$.

3.2.1. Theoretical Comparisons

The likelihood function of θ based on the sample is

$$L(\theta) = \frac{\exp(-\sum_{i=1}^{n} x_i)}{\left(e^{-\theta} - e^{-\theta^2}\right)^n} I_{[\theta,\theta^2]}(t_1) I_{[\theta,\theta^2]}(t_n) \propto \frac{I_{[\sqrt{t_n},t_1]}(\theta)}{\left(e^{-\theta} - e^{-\theta^2}\right)^n}.$$

The MLE. In this example, the MLE of θ is unique and given by

$$\hat{\theta}_{\mathrm{M}} = t_1\Delta + \sqrt{t_n}(1 - \Delta), \tag{16}$$

where $\Delta = 1$ if $e^{-\sqrt{t_n}} - e^{-t_n} > e^{-\sqrt{t_1}} - e^{-t_1}$ and 0 otherwise. The MLE $\hat{\theta}_{\mathrm{M}}$ in (16) has an asymptotic distribution with the same form as that in (6) with $\lambda_1 = 1 - e^{\theta-\theta^2}$, $\lambda_2 = \{e^{\theta^2-\theta} - 1\}/\{2\theta\}$ if $2\theta e^{-\theta^2} \neq e^{-\theta}$. Based on (6) and (16), the shortest confidence interval of θ with nominal coverage probability α is

$$\mathrm{CI}_{\mathrm{M}} = \begin{cases} \left[t_1 + \frac{\log(1-\alpha)\hat{\lambda}_1}{n}, \quad t_1\right], & \text{if } \Delta = 1, \\ \left[\sqrt{t_n}, \quad \sqrt{t_n} - \frac{\log(1-\alpha)\hat{\lambda}_2}{n}\right], & \text{if } \Delta = 0, \end{cases}$$

where $\hat{\lambda}_1 = 1 - e^{t_1 - t_1^2}$ and $\hat{\lambda}_2 = \{e^{t_n - \sqrt{t_n}} - 1\}/\{2\sqrt{t_n}\}$.

This is an example that the MLE is asymptotically biased almost everywhere (the asymptotic bias is λ_1 or $-\lambda_2$). If θ is large, $e^{-\theta^2}$ is much smaller than $e^{-\theta}$. Then λ_1 is close to 1 and λ_2 is much greater than λ_1. So the asymptotic absolute bias and the asymptotic variance of the MLE both approach 1 as θ increases.

Bayesian Estimation. The reference prior of θ for model (15) is the same as equation (4) with $b_1 = |1 - 2\theta e^{\theta-\theta^2}|$, $b_2 = |1 - 1/\{2\theta e^{\theta-\theta^2}\}|$ and $\lambda = \{e^{-\theta} - e^{-\theta^2}\}/\{2\theta e^{-\theta^2} - e^{-\theta}\}$ if $2\theta e^{-\theta^2} \neq e^{-\theta}$.

The posterior density of θ given (t_1, t_n) under a prior $\pi(\theta)$ is

$$\pi_{\mathrm{B}}(\theta \mid t_1, \sqrt{t_n}) \propto \frac{\pi(\theta)}{\left(e^{-\theta} - e^{-\theta^2}\right)^n} I_{[\sqrt{t_n}, t_1]}(\theta). \tag{17}$$

Under the square error loss, the Bayes estimator of θ under a prior $\pi(\theta)$ is

$$\hat{\theta}_B = \int_{\sqrt{t_n}}^{t_1} \frac{s\pi(s)}{\left(e^{-s} - e^{-s^2}\right)^n} \mathrm{d}s \bigg/ \int_{\sqrt{t_n}}^{t_1} \frac{\pi(s)}{\left(e^{-s} - e^{-s^2}\right)^n} \mathrm{d}s, \tag{18}$$

of which the asymptotic distribution Λ takes the same form as equation (7).

For large θ, λ is close to -1 and $e^{\{\lambda_2 Z_2\}/\lambda}$ is close to 0, so the asymptotic distribution Λ of the Bayes estimator is close to that of $Z_1 - 1$, a random variable with mean 0 and variance 1. This indicates, asymptotically, the Bayes estimator is approximately twice as efficient as the MLE for large θ.

3.2.2. Numerical Comparisons

All the results reported in Figures .3-.6 are based on 1000 repetitions of the simulation.

Figure .3 exhibits the RMSE of $\hat{\theta}_{\mathrm{M}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ compared to $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$, where $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ is the Bayes estimator in (18) under the reference prior and $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ is the Bayes estimator under the constant prior. From (a) of Figure .3, it is seen that the RMSE are always greater than 2, which means that $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ is as at least twice efficient as $\hat{\theta}_{\mathrm{M}}$. The results are especially significant when θ is between 1 and 2. From (b) of Figure .3, the RMSE stay closely to 1 when $n = 100$, indicating the equivalence of $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ for a large sample size. For small sample sizes, neither $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ or $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ dominates each other. But the advantage of $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ is more significant than that of $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$.

Figure .4 presents the asymptotic MSEs and variances of Bayes estimators and the MLE. One can see that the asymptotic MSEs of the MLE are at least twice as that of the Bayes estimators.

Figure .5 gives the estimated densities of each point estimators from the simulated samples of size 5. When $\theta = 1.45$, clearly the distribution of $\hat{\theta}_{\mathrm{M}}$ is skewed and biased whereas the distributions of $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ are symmetric of the true value. When $\theta = 1.718$, a value such that λ_1 and λ_2 are close, the densities of all estimators are approximately symmetric of the true value. However, the density of $\hat{\theta}_{\mathrm{M}}$ is not uni-model and has a much bigger variance than the other two. When $\theta = 2.5$, $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ are also skewed and the bias dominates the difference of the performance between Bayes estimators and the MLE.

Figure .6 is about the frequentist coverage probabilities of Bayesian credible intervals, which are lower one sided intervals calculated from the posterior distribution in (17) with $\pi(\theta)$ being replaced by either the reference prior ($\mathrm{CI}_{\mathrm{B}}^{\mathrm{R}}$) or the constant prior ($\mathrm{CI}_{\mathrm{B}}^{\mathrm{C}}$). It shows that when $n = 5$, all the three intervals give desirable outcomes. $\mathrm{CI}_{\mathrm{B}}^{\mathrm{R}}$ evidently outperforms $\mathrm{CI}_{\mathrm{B}}^{\mathrm{C}}$ when $n = 1, 2$, and the MLE is not capable of yielding satisfactory outcomes in this scenario.

4. Summary

In this paper, the reference prior is developed for a truncated model and the asymptotic properties of Bayes estimators are obtained. Numerical studies are carried out to evaluate the performance of both the Bayesian and frequentist approaches. According to both the theoretical and the numerical outcomes, Bayes estimators are superior alternatives to the frequentist counterparts, the MLE, and the reference prior is preferable over the constant prior for Bayesian approach. However the model of interest in this paper requires that $g(\cdot)$ is independent of θ. In practice, $g(\cdot)$ often depends on θ, and then (t_1, t_n) may not be a sufficient statistic of θ. Reference priors for this case are tremendously more challenging to find and this is a topic worthwhile for future investigation.

Acknowledgements

This research was partially supported by NSF grants DMS-1007874 and SES-1024080. We are very grateful to two referees for constructive suggestions and comments which have greatly improved the presentation of this paper. We would also like to thank an associate editor for valuable comments.

Appendix A. Proof of Theorem 1

Let $x_1, ..., x_n$ be a simple random sample from model (1). (t_1, t_n) is a sufficient statistic of θ whose joint density is

$$f(t_1, t_n \mid \theta) = \frac{n(n-1)g(t_1)g(t_n)\big[G(t_n) - G(t_1)\big]^{n-2}}{\left[G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)\right]^n}, \quad a_1(\theta) \le t_1 < t_n \le a_2(\theta).$$

Denote by $y_1 = n[a_1^{-1}(t_1) - \theta]$ and $y_2 = n[\theta - a_2^{-1}(t_n)]$. Then

$$\begin{aligned}
f(y_1, y_2 \mid \theta) =& \frac{n-1}{n} \frac{a_1'(\theta + \frac{y_1}{n})a_2'(\theta - \frac{y_2}{n})g\big(a_1(\theta + \frac{y_1}{n})\big)g\big(a_2(\theta - \frac{y_2}{n})\big)}{\left[G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)\right]^n} \\
&\times \left[G\left(a_2(\theta - \frac{y_2}{n})\right) - G\left(a_1(\theta + \frac{y_1}{n})\right)\right]^{n-2} \\
=& \frac{n-1}{n} \frac{a_1'(\theta + \frac{y_1}{n})a_2'(\theta - \frac{y_2}{n})g\big(a_1(\theta + \frac{y_1}{n})\big)g\big(a_2(\theta - \frac{y_2}{n})\big)}{\left[G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)\right]^2} \\
&\times \left\{\frac{G\left(a_2(\theta - \frac{y_2}{n})\right) - G\left(a_1(\theta + \frac{y_1}{n})\right)}{G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)}\right\}^{n-2} \\
=& \frac{1}{\lambda_1\lambda_2}\left\{1 - \frac{1}{n}\frac{g\big(a_1(\theta)\big)a_1'(\theta)y_1 + g\big(a_2(\theta)\big)a_2'(\theta)y_2}{G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)} + o(\frac{1}{n})\right\}^{n-2}\big[1 + o(1)\big] \\
\to& \frac{1}{\lambda_1\lambda_2}\exp\left(-\frac{g\big(a_1(\theta)\big)a_1'(\theta)y_1 + g\big(a_2(\theta)\big)a_2'(\theta)y_2}{G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)}\right) \\
=& \frac{1}{\lambda_1}\exp(-\frac{y_1}{\lambda_1}) \times \frac{1}{\lambda_2}\exp(-\frac{y_2}{\lambda_2}) \equiv f^*(y_1, y_2 \mid \theta).
\end{aligned}$$

This means y_1 and y_2 are, asymptotically, independent exponential variables with mean λ_1 and λ_2, respectively. Choosing a prior of θ as $\pi^* = 1$, the posterior density given (t_1, t_n) is

$$\pi^*(\theta \mid t_1, t_n) = \frac{1}{\left[G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)\right]^n m_n(t_1, t_n)}, \quad \theta \in \left[a_2^{-1}(t_n), a_1^{-1}(t_1)\right],$$

where

$$m_n(t_1,t_n) = \int_{a_2^{-1}(t_n)}^{a_1^{-1}(t_1)} \frac{1}{\left[G\left(a_2(s)\right) - G\left(a_1(s)\right)\right]^n} \mathrm{d}s$$
$$= \frac{1}{n}\int_{-y_2}^{y_1} \frac{1}{\left[G\left(a_2(\theta+\frac{v}{n})\right) - G\left(a_1(\theta+\frac{v}{n})\right)\right]^n} \mathrm{d}v.$$

The last equality holds from the transformation $s = \theta + v/n$. Thus for fixed y_1 and y_2,

$$nm_n(t_1,t_n)\left[G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)\right]^n = \int_{-y_2}^{y_1} \frac{\left[G\left(a_2(\theta)\right) - G\left(a_1(\theta)\right)\right]^n}{\left[G\left(a_2(\theta+\frac{v}{n})\right) - G\left(a_1(\theta+\frac{v}{n})\right)\right]^n} \mathrm{d}v$$

$$\longrightarrow \int_{-y_2}^{y_1} \exp\Big(-[\frac{1}{\lambda_2}-\frac{1}{\lambda_1}]v\Big)\mathrm{d}v = \begin{cases} |\lambda|\exp(\frac{y_2}{|\lambda|})\Big\{1-\exp\big[-\frac{y_1+y_2}{|\lambda|}\big]\Big\}, & \text{if } \lambda_1 > \lambda_2, \\ |\lambda|\exp(\frac{y_1}{|\lambda|})\Big\{1-\exp\big[-\frac{y_1+y_2}{|\lambda|}\big]\Big\}, & \text{if } \lambda_1 < \lambda_2, \\ y_1+y_2, & \text{if } \lambda_1 = \lambda_2, \end{cases}$$

$$= |\lambda|\exp\left\{\frac{y_1 I_{(\lambda_1<\lambda_2)} + y_2 I_{(\lambda_1>\lambda_2)}}{|\lambda|}\right\}\left\{1-\exp\big[-\frac{y_1+y_2}{|\lambda|}\big]\right\} I_{(\lambda_1\neq\lambda_2)} + (y_1+y_2)I_{(\lambda_1=\lambda_2)}.$$

So for fixed θ,

$$-\int \log\big[\pi^*(\theta \mid t_1,t_n)\big] f(t_1,t_n\mid\theta)\mathrm{d}t_1\mathrm{d}t_n + \log(n)$$
$$\longrightarrow\Big\{\log(|\lambda|) + \frac{\mathbf{E}\big[y_1 I_{(\lambda_1<\lambda_2)} + y_2 I_{(\lambda_1>\lambda_2)}\big]}{|\lambda|} + \mathbf{E}\log\Big\{1-\exp\Big(-\frac{y_1+y_2}{|\lambda|}\Big)\Big\}\Big\} I_{(\lambda_1\neq\lambda_2)}$$
$$+ \mathbf{E}\log(y_1+y_2)I_{(\lambda_1=\lambda_2)}$$
$$=\Big\{\log(|\lambda|) + \frac{\lambda_1 I_{(\lambda_1<\lambda_2)} + \lambda_2 I_{(\lambda_1>\lambda_2)}}{|\lambda|} - \sum_{i=1}^{\infty}\frac{1}{i}\mathbf{E}\Big\{\exp\Big[-\frac{i(y_1+y_2)}{|\lambda|}\Big]\Big\}\Big\} I_{(\lambda_1\neq\lambda_2)}$$
$$+ \Big[\log(\lambda_1) + \mathbf{E}\log((y_1+y_2)/\lambda_1)\Big] I_{(\lambda_1=\lambda_2)}$$
$$=\Big\{\log(|\lambda|) + \frac{\lambda_1\wedge\lambda_2}{|\lambda|} - \sum_{i=1}^{\infty}\frac{1}{i}\mathbf{E}\exp(-\frac{iy_1}{|\lambda|})\mathbf{E}\exp(-\frac{iy_2}{|\lambda|})\Big\} I_{(\lambda_1\neq\lambda_2)} + \big[\log(\lambda_1)+C\big] I_{(\lambda_1=\lambda_2)}$$
$$=\Big\{\log(|\lambda|) + b_1\wedge b_2 - \sum_{i=1}^{\infty}\frac{1}{i(1+b_1 i)(1+b_2 i)}\Big\} I_{(\lambda_1\neq\lambda_2)} + \big[\log(\lambda_1)+C\big] I_{(\lambda_1=\lambda_2)}$$
$$=\Big\{\log(|\lambda|) + b_1\wedge b_2 - \gamma - \frac{1}{b_1-b_2}\Big\{b_1\psi(\frac{1}{b_1}+1) - b_2\psi(\frac{1}{b_2}+1)\Big\}\Big\} I_{(\lambda_1\neq\lambda_2)}$$
$$+ \big[\log(\lambda_1)+C\big] I_{(\lambda_1=\lambda_2)},$$

where the expectations are taken with respect to y_1 and y_2 under the density of $f^*(y_1,y_2 \mid \theta)$, $C = \int_0^\infty\int_0^\infty \log(z_1+z_2)e^{-z_1-z_2}\mathrm{d}z_1\mathrm{d}z_2 = 1-\gamma$, and $\gamma \approx 0.5772$ is the Euler-Mascheroni constant. Using the fact that $\psi(z+1) = \psi(z) + 1/z$,

$$\int \log\big[\pi^*(\theta) \mid t_1,t_n\big] f(t_1,t_n\mid\theta)\mathrm{d}t_1\mathrm{d}t_n - \log(n)$$
$$= \Big\{-\log(|\lambda|) + b_1\vee b_2 + 1 + \frac{1}{b_1-b_2}\Big\{b_1\psi(\frac{1}{b_1}) - b_2\psi(\frac{1}{b_2})\Big\}\Big\} I_{(\lambda_1\neq\lambda_2)} - \log(\lambda_1)I_{(\lambda_1=\lambda_2)} - C.$$

Finally, the reference prior follows from Theorem 7 of Berger *et al.* (2009).

Appendix B. Proof of Theorem 2

The case when $\lambda_1 = \lambda_2$ is straightforward so we prove only when $\lambda_1 \neq \lambda_2$.

$$n\left(\hat{\theta}_B - \theta\right) = \frac{\int_{a_2^{-1}(t_n)}^{a_1^{-1}(t_1)} \frac{n(s-\theta)\pi(s)}{[G(a_2(s))-G(a_1(s))]^n} \mathrm{d}s}{\int_{a_2^{-1}(t_n)}^{a_1^{-1}(t_1)} \frac{\pi(s)}{[G(a_2(s))-G(a_1(s))]^n} \mathrm{d}s} = \frac{\int_{-y_2}^{y_1} \frac{v\pi(\theta+\frac{v}{n})}{\left[G\left(a_2(\theta+\frac{v}{n})\right)-G\left(a_1(\theta+\frac{v}{n})\right)\right]^n} \mathrm{d}v}{\int_{-y_2}^{y_1} \frac{\pi(\theta+\frac{v}{n})}{\left[G\left(a_2(\theta+\frac{v}{n})\right)-G\left(a_1(\theta+\frac{v}{n})\right)\right]^n} \mathrm{d}v}$$

$$= \frac{\int_{-y_2}^{y_1} v\left[\frac{G(a_2(\theta))-G(a_1(\theta))}{G\left(a_2(\theta+\frac{v}{n})\right)-G\left(a_1(\theta+\frac{v}{n})\right)}\right]^n \mathrm{d}v}{\int_{-y_2}^{y_1} \left[\frac{G(a_2(\theta))-G(a_1(\theta))}{G\left(a_2(\theta+\frac{v}{n})\right)-G\left(a_1(\theta+\frac{v}{n})\right)}\right]^n \mathrm{d}v} + o_P(1) \xrightarrow{d} \frac{\int_{-\lambda_2 Z_2}^{\lambda_1 Z_1} s e^{-\frac{s}{\lambda}} \mathrm{d}s}{\int_{-\lambda_2 Z_2}^{\lambda_1 Z_1} e^{-\frac{s}{\lambda}} \mathrm{d}s} = \Lambda.$$

The convergence in the last line follows from the continuous mapping theorem and Slutsky's theorem. For the asymptotic unbiasedness, notice

$$\mathbf{E}\Lambda = \lambda \int_0^\infty \int_0^\infty \frac{(1-\frac{\lambda_2 z_2}{\lambda})e^{\frac{\lambda_2 z_2}{\lambda}} - (\frac{\lambda_1 z_1}{\lambda}+1)e^{-\frac{\lambda_1 z_1}{\lambda}}}{e^{\frac{\lambda_2 z_2}{\lambda}} - e^{\frac{-\lambda_1 z_1}{\lambda}}} e^{-z_1-z_2} \mathrm{d}z_1 \mathrm{d}z_2.$$

Let $v_1 = \lambda_1 z_1/\lambda$, $v_2 = \lambda_2 z_2/\lambda$ and $r = \lambda_1/\lambda$. Then from Fubini's theorem,

$$\begin{aligned}
\mathbf{E}\Lambda =& \frac{1+r}{r^2} \int_0^\infty \int_0^\infty \frac{(1-v_2)-(1+v_1)e^{-v_1-v_2}}{1-e^{-v_1-v_2}} e^{-\frac{v_1}{r}-\frac{(1+r)v_2}{r}} \mathrm{d}v_1 \mathrm{d}v_2 \\
=& \frac{1+r}{r^2} \int_0^\infty \int_0^\infty [(1-v_2)-(1+v_1)e^{-v_1-v_2}] e^{-\frac{v_1}{r}-\frac{(1+r)v_2}{r}} \sum_{i=0}^\infty e^{-iv_1-iv_2} \mathrm{d}v_1 \mathrm{d}v_2 \\
=& \frac{1+r}{r^2} \sum_{i=0}^\infty \int_0^\infty \int_0^\infty [(1-v_2)-(1+v_1)e^{-v_1-v_2}] e^{-\frac{v_1}{r}-\frac{(1+r)v_2}{r}} e^{-iv_1-iv_2} \mathrm{d}v_1 \mathrm{d}v_2 \\
=& \frac{1+r}{r^2} \sum_{i=0}^\infty \left\{ \left[\frac{1}{i+1+\frac{1}{r}} - \frac{1}{\left(i+1+\frac{1}{r}\right)^2}\right] \frac{1}{i+\frac{1}{r}} - \left[\frac{1}{i+1+\frac{1}{r}} + \frac{1}{\left(i+1+\frac{1}{r}\right)^2}\right] \frac{1}{i+2+\frac{1}{r}} \right\} \\
=& 0.
\end{aligned}$$

This finishes the proof.

References

Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *Journal of the American Statistical Association* **84**, 200–207.

Berger, J. O. and Bernardo, J. M. (1992a). On the development of reference priors. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds., *Bayesian statistics*, vol. 4, 35–60. Oxford Univ. Press, New York.

Berger, J. O. and Bernardo, J. M. (1992b). Ordered group reference priors with application to the multinomial problem. *Biometrika* **79**, 25–37.

Berger, J. O. and Bernardo, J. M. (1992c). Reference priors in a variance components problem. In P. K. Goel and N. S. Iyengar, eds., *Bayesian analysis in statistics and econometrics*, vol. 75 of *Lecture Notes in Statist.*, 177–194. Springer, New York.

Berger, J. O., Bernardo, J. M., and Sun, D. (2009). The formal definition of reference priors. *The Annals of Statistics* **37**, 905–938.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)* **41**, 113–147.

Bernardo, J. M. (2005). Reference analysis. In D. K. Dey and C. R. Rao, eds., *Handbook of Statistics*, vol. 25, 17–90. North-Holland, Amsterdam.

Ghosal, S. (1997). Reference priors in multiparameter nonregular cases. *Test* **6**, 159–186.

Ghosal, S. and Samanta, T. (1997). Expansion of bayes risk for entropy loss and reference prior in nonregular cases. *Statistcs & Decisions* **15**, 129–140.

Hall, P. and Wang, J. Z. (2005). Bayesian likelihood methods for estimating the end point of a distribution. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **67**, 717–729.

Hill, B. M. (1963). The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic. *Journal of the American Statistical Association* **58**, 72–84.

Hirano, K. and Porter, J. R. (2003). Asymptotic efficiency in parametric structural models with parameter-dependent support. *Econometrica* **71**, 1307–1338.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **186**, 453–461.

Jeffreys, H. (1961). *Theory of probability.* Third edition. Clarendon Press, Oxford.

Sun, D. and Berger, J. O. (1998). Reference priors with partial information. *Biometrika* **85**, 55–71.

Sun, D. and Ye, K. (1995). Reference prior Bayesian analysis for normal mean products. *Journal of the American Statistical Association* **90**, 589–597.

Yang, R. and Berger, J. O. (1994). Estimation of a covariance matrix using the reference prior. *The Annals of Statistics* **22**, 1195–1211.

Ye, K. and Berger, J. O. (1991). Noninformative priors for inferences in exponential regression models. *Biometrika* **78**, 645–656.

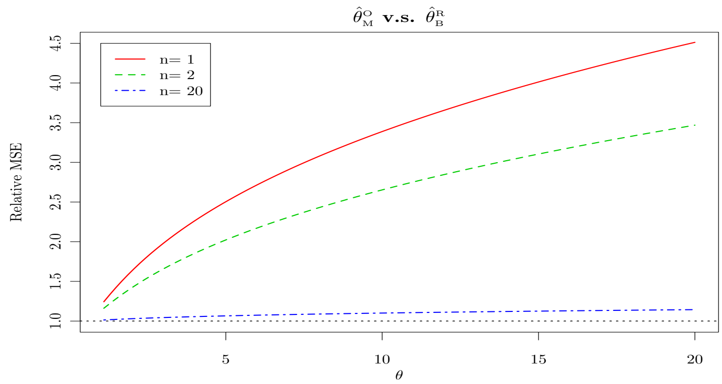

Figure .1: Relative MSEs of $\hat{\theta}_{\mathrm{M}}^{\mathrm{O}}$ compared to $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ as functions of θ for different sample size n.

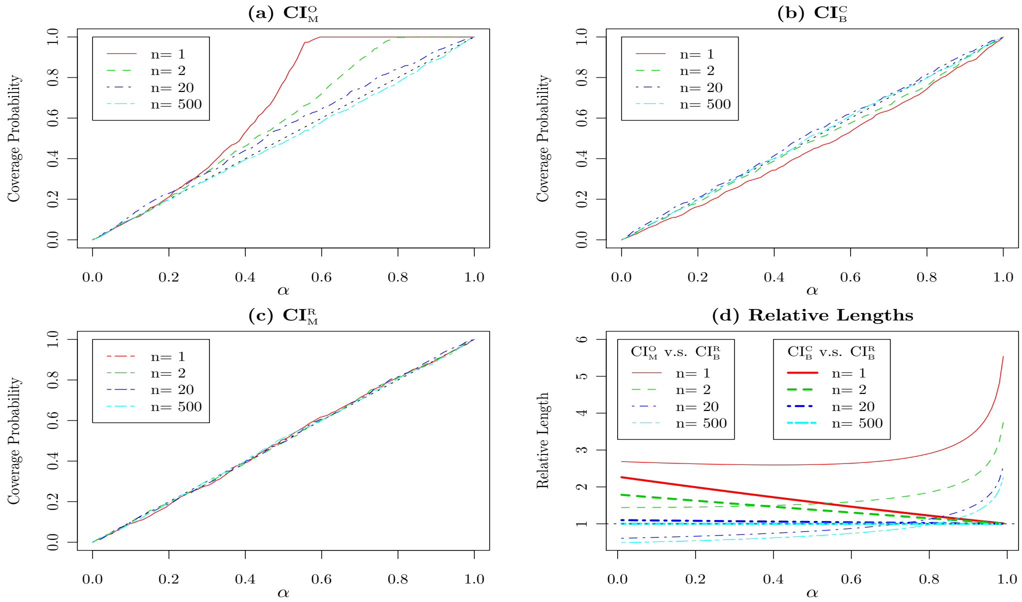

Figure .2: Coverage probabilities of $\mathrm{CI}_{\mathrm{M}}^{\mathrm{O}}$, $\mathrm{CI}_{\mathrm{B}}^{\mathrm{C}}$ and $\mathrm{CI}_{\mathrm{B}}^{\mathrm{R}}$ and their relative lengths at the base of $\mathrm{CI}_{\mathrm{B}}^{\mathrm{R}}$ as functions of α for different sample size n when $\theta = 3$.

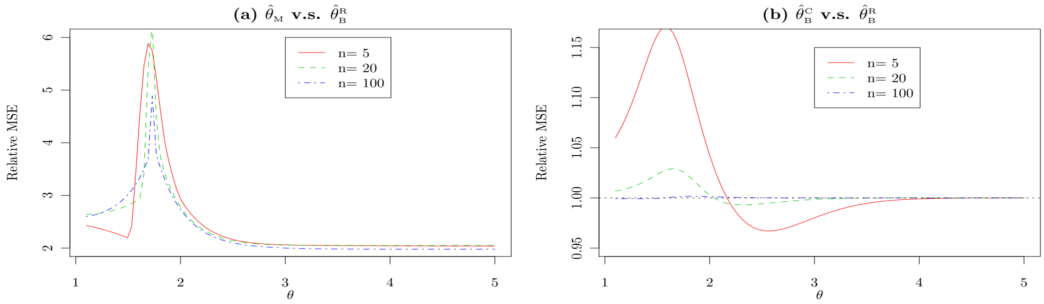

Figure .3: Relative MSEs of $\hat{\theta}_{\mathrm{M}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ compared to $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ as functions of θ for different sample size n.

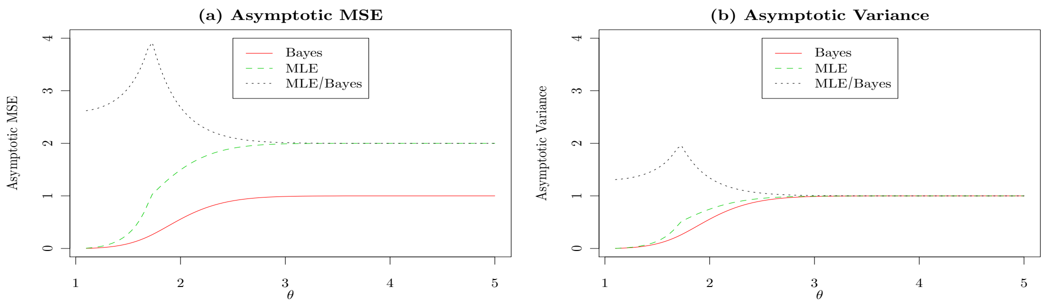

Figure .4: Asymptotic MSEs and variances as functions of θ (The red lines in (a) and (b) represent the MSE and variance, respectively, of the asymptotic distribution of Bayes estimators, whereas green lines represent those quantities for the MLE. Dotted lines represent the ratio of each quantity of the MLE to that of Bayes estimators).

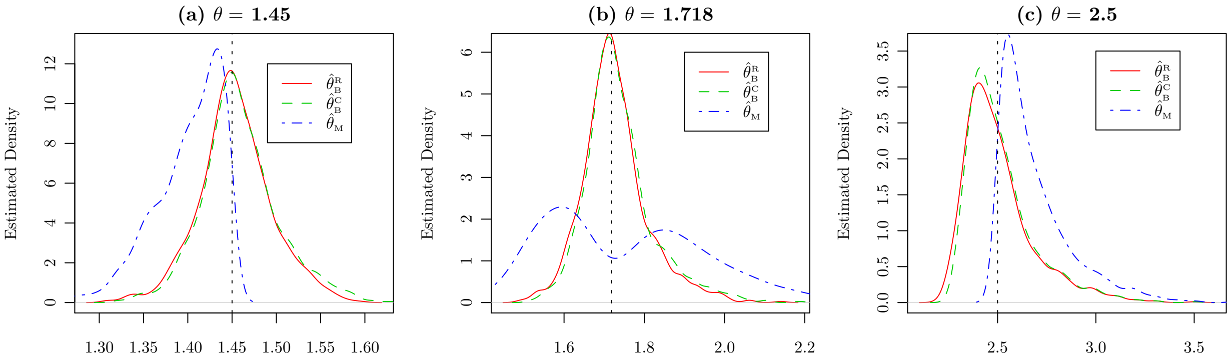

Figure .5: Estimated densities of $\hat{\theta}_{\mathrm{M}}$, $\hat{\theta}_{\mathrm{B}}^{\mathrm{C}}$ and $\hat{\theta}_{\mathrm{B}}^{\mathrm{R}}$ for different θ when $n = 5$.

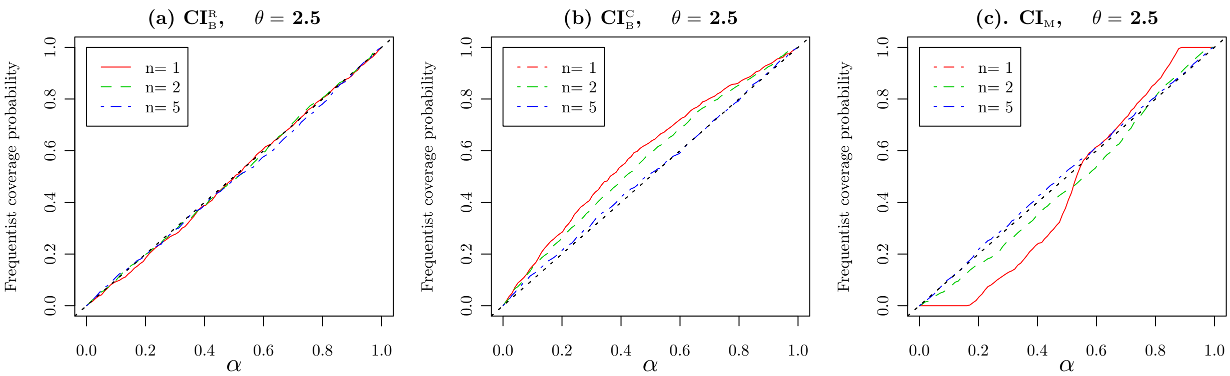

Figure .6: Coverage probabilities of $\mathrm{CI}_{\mathrm{M}}^{\mathrm{O}}$, $\mathrm{CI}_{\mathrm{B}}^{\mathrm{C}}$ and $\mathrm{CI}_{\mathrm{B}}^{\mathrm{R}}$ as functions of α for different sample size n and θ.